

Title	戦略の進化 繰り返し囚人のディレンマ・ゲームのコンピュータ・シミュレーション
Author(s)	高増, 明
Citation	経済論叢 (1999), 164(5): 100-120
Issue Date	1999-11
URL	http://dx.doi.org/10.14989/45314
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

經濟論叢

第164巻 第5号

瀬地山 敏教授記念號

献 辞.....	西 村 周 三	
ミクロ・マクロ・ループについて.....	塩 沢 由 典	1
進化経済学と複雑系.....	有 賀 裕 二	74
戦 略 の 進 化.....	高 増 明	100
不確実性下の意思決定理論.....	竹 治 康 公	121
非平衡非線形経済システム試論.....	吉 田 和 男	145
H.J. ダヴェンポートの貨幣的マクロ経済理論.....	小 島 專 孝	162

瀬地山 敏 教授 略歴・著作目録

平成11年11月

京都大學經濟學會

戦 略 の 進 化

——繰り返し囚人のディレンマ・ゲームの
コンピュータ・シミュレーション——

高 増 明

I は じ め に

経済学は、これまで「合理的選択」と「均衡」をもっとも基本的な概念として、その理論体系を発展させてきた。経済主体は、無数の選択肢のなかから最適な行動を選択し、各経済主体の最適行動の結果として経済の均衡が達成される。経済学の分析は、体系の外部から与えたパラメータを変化させたときに、各経済主体の行動（選択）と均衡がどのように変化するかを考察するものだと単純化しても、それほど間違っていないだろう。

もちろん、このような分析は非現実的なものである。企業や消費者などの経済主体は現実にはそれほど合理的に行動しているとは考えられないし、現実の経済の価格や生産量が経済モデルの均衡として説明できるかについても問題がある。したがって、これまでも多くの経済学者がこうした分析方法を批判してきたのだが¹⁾、批判者がそれに替わる論理的で説得的な分析ツールを提示することができなかったのも事実である。

しかしながら、最近の複雑系や人工生命に関する理論やコンピュータ・シミュレーションの発展、経済学の分野でそれに触発されて誕生した進化経済

1) Robinson [1971] は、主流派経済学の均衡概念を痛烈に批判している。また Sraffa [1960] のように経済主体の行動を考慮しないで、経済の再生産構造に注目するアプローチも存在する。しかし、経済が人間行動によって産み出されるものである以上、人間の行動についての理論なしに経済学を発展させていくには自ずと限界があるだろう。

学²⁾などの分析は、私たちに、これまでの経済学とは違う新たな分析方法の可能性を示している。

この論文でも、そのようなアプローチを使って、最適化行動をとらない経済主体の相互作用のなかで、どのようにして戦略が進化し、社会秩序が形成されていくのかが分析される。分析に使用されるモデルは、「繰り返し囚人のディレンマゲーム」と呼ばれるものである。このゲームでは、二人のプレイヤーが対戦するが、二人のプレイヤーが短期的に利得を最大にするように利己的な選択を行うと、結局はどちらにとっても不利な結果が生じてしまう状況が想定されている。そのような状況で、どのようにしてプレイヤーの協調が実現されるのかが、様々なアプローチによって研究されている。

最近では、とくに Axelrod [1980]³⁾ によって開始されたコンピュータ・シミュレーションが注目されている。これは、多様な戦略が参加して、このゲームを対戦するときに、どの戦略が勝つかをコンピュータ・プログラムのリーグ戦として行うものである。

ただし、社会科学では、より強い戦略を探すことよりも、様々な戦略が対戦し、新しい戦略が生まれ、弱い戦略が減んでいくなかで、どのようにして社会秩序が形成されていくのかに興味の中心がある。

そのようなアプローチの代表的なものとしては、Lindgren [1991], Lindgren and Nordahl [1993], [1996] がある。また高増・服部 [1999a] は、Lindgren and Nordahl のアプローチをより厳密なかたちで再定式化し、メモリー 1 の 4

2) 瀬地山敏教授が会長を務める進化経済学会は、日本におけるこうしたアプローチの中心となる存在である。ただし、学会の追求する方向性は正しいとしても、こうしたアプローチがあいまいな言葉による説明、厳密ではない「理論」に流れてしまっただけでは、これまでの経済学批判と同じような運命をたどることになるだろう。主流派経済学の理論的發展を踏まえたうえで、それを乗り越えるだけの厳密な理論体系を發展させていく必要があるのではないだろうか。

3) Axelrod が行った繰り返し囚人のディレンマについてのコンピュータ・プログラムの第1回目のトーナメントには15の戦略が参加し、しっぺ返し (Tit For Tat) 戦略が勝利した。その結果の詳細を公表してから行った第2回のトーナメントには、62の戦略が参加したが、そこでも、やはり TFT が勝利を収めた。トーナメントの結果とその分析については、Axelrod [1980a], [1980b], [1984] Chap. 2 に掲載されている。また繰り返し囚人のディレンマのコンピュータ・トーナメントについては、Axelrod [1987], [1997], Poundstone [1992] も参照してもらいたい。

種類の戦略がランダムに出会って戦うなかで、どの戦略が生き残るのかを利得行列や初期の個体数を変化させて検討している。さらに高増・服部 [1999b] では、プレイヤーがミスする可能性（ノイズ）を考慮するときに、シミュレーションの結果がどのように変わるのか、またメモリーのサイズを1から2に増加したときに、どのような結果が得られるのかを検討している。

この論文では、高増・服部 [1999a], [1999b] の分析に、戦略の進化を導入する。そして、それが繰り返し囚人のディレンマ・ゲームのコンピュータ・シミュレーションにどのような影響をもたらすのかを可能な限り厳密に検討する。

論文の構成を簡単に紹介しておこう。まず第Ⅱ節では、繰り返し囚人のディレンマゲームの構造とそれを様々なアプローチがどのように取り扱っているのかについて簡単に紹介する。第Ⅲ節では、繰り返し囚人のディレンマゲームで戦略の進化をどのように取り扱うのかについて議論する。第Ⅳ節では、戦略のメモリーサイズが1, 2, 3のケースについてコンピュータ・シミュレーションを行う。最後に、これまでに得られた結果が再び吟味される。

Ⅱ 繰り返し囚人のディレンマゲーム

1 囚人のディレンマゲーム

囚人のディレンマとは、つぎのような状況のことを指す。

裁判をひかえた二人の囚人がいる。二人は共同で重大な犯罪を行ったことがわかっているが、検察側はそれを立証するのに十分な証拠をもっていない。ただし、軽い犯罪についての証拠は揃っている。そこで、囚人はそれぞれ別々に尋問されて自白を促されている。二人がともに自白すれば、かれらは、懲役10年となる。両者が完全に黙秘すれば、軽い犯罪による2年の刑だけですむ。一人が自白し、もう一人が黙秘した場合には、自白した者は釈放され、他方は20年の刑になる。このような状況が囚人のディレンマと呼ばれる状況である。

このとき二人の囚人はどのような行動をとるだろうか。二人の行動、すなわ

ち黙秘、自白と二人の刑期の関係は以下の表1のようにになっている。表1でカッコの最初の数字は囚人1の刑期を2番目の数字は囚人2の刑期を表している。

表1 囚人のディレンマ

囚人2

囚人1		黙 秘	自 白
	黙 秘	(2, 2)	(20, 0)
	自 白	(0, 20)	(10, 10)

囚人1がどのように行動するのかを考えてみよう。囚人1は、囚人2が黙秘するときには、黙秘すれば2年の刑、自白すれば釈放されるから自白するのが有利である。一方、囚人2が自白するときには、黙秘すれば20年の刑、自白すれば10年の刑なので、やはり自白するのが有利である。このように囚人1にとっては囚人2がどのような行動をとろうと自白するのが有利になる⁴⁾。このことは、囚人2にとっても同様であるから、結局どちらの囚人も自白し、結果として、どちらも10年の刑に服することになる。しかしながら、二人が単純に相手のことを考えて黙秘という行動を選択すれば、どちらも2年の刑ですむわけである。このように二人が合理的に行動した結果、どちらにとっても不利な結果になってしまう状況が囚人のディレンマである。

この囚人のディレンマが注目されるのは、このような状況が、社会に広範に存在しているからである。プレイヤーを囚人のかわりに企業や様々な団体、国家などに置き換え、二つの選択肢を協調的な戦略、裏切りのな戦略に置き換えると、このゲームと同じ構造をもつ現実の状況をいくらかでも思い浮かべることができるだろう。

状況をより一般的なものにするために、自白と黙秘という二つの戦略を裏切り (Defect) と協調 (Cooperation) に、また刑期を表2のような利得行列に変更してやろう。

4) このような戦略を支配戦略 (dominant strategy) という。

表2 囚人のディレンマの利得行列による定義(数値を与えたもの)
プレイヤー2

プレイヤー1		協 調 (C)	裏 切 り (D)
	協 調 (C)	(3, 3)	(0, 5)
	裏 切 り (D)	(5, 0)	(1, 1)

このゲームでは、どちらのプレイヤーも裏切り (D) を選択することが支配戦略であるが、表2ではどちらも1の利得しか得ることはできない。一方、どちらも協調を選択しておけば3の利得を得ることができる。

2 繰り返し囚人のディレンマ・ゲーム

ゲームが1回かぎりのときには、どちらのプレイヤーにとっても裏切りが支配戦略であり、したがって二人のプレイヤーが裏切りを選択し、どちらも1の利得を得る結果が唯一のナッシュ均衡になる。しかし、このゲームが長期間繰り返し行われるとしたら、どうなるだろうか。裏切り続ければ、ずっと1の利得しか得られないのに対して、協調を続ければ3の利得が得られる。したがって、囚人のディレンマの状況が繰り返し行われる「繰り返し囚人のディレンマゲーム」では、協調が最適な戦略となる可能性があることが予想できる。

ゲーム理論では、個々のプレイヤーが将来のことを重要視するならば、すなわち将来の利得の割引率が十分に1に近いときには、どちらも「協調」を選択するという「協調解」を含む様々な解が、繰り返し囚人のディレンマゲームのナッシュ均衡となるという「フォーク (folk) 定理」⁵⁾ が証明されている。

しかし、このフォーク定理や一般にゲーム理論の分析にも様々な問題点が指摘されている⁶⁾。まずプレイヤーについては、無数の選択肢のなかから最適な

5) folk 定理については、たとえば Fudenberg and Tirole [1992], pp. 150-155, Binmore [1992], pp. 369-377 を参照せよ。この定理が folk 定理と呼ばれているのは、この定理が古くから知られていて、しかも誰の貢献によるものか確かではない、すなわち folklore だからである。

6) ゲーム理論、とくにナッシュ均衡の問題点については、Kreps [1990] の第5章を参照してもらいたい。

ものを見つけ選択することができるという神のような合理性が仮定されている。またゲーム理論のもっとも基本的な均衡概念はナッシュ均衡であるが、ナッシュ均衡は一般に複数個存在し、そのうちどれが現実を選択されるかは自明ではないという問題もある。

一方、このようなゲーム理論とは対照的なアプローチとして、生物学の進化ゲーム⁷⁾がある。進化ゲームでは、プレイヤーは、自分の戦略を変化させない。しかし、相互に対戦するなかで、利得が大きい戦略はその数を増やし、少ない戦略は数を減らし、その結果として社会の構造が変化していくと考える。その意味で、ゲーム理論の非現実性という問題点は免れている。したがって、ここでも、進化ゲームのアプローチを使って、様々な戦略がランダムに出会って対戦していくなかで、高得点を得た戦略の個体数が増え、得点の低い戦略の個体数は減っていく、としよう。

3 戦略とメモリー・サイズ

各プレイヤーが、その対戦相手の過去に取った戦略をどれだけ記憶しているのかをメモリーサイズと呼ぶことにする⁸⁾。メモリーサイズが1のときには、各プレイヤーは、対戦相手が前回に選択した手だけを記憶している。メモリーサイズが1だとして、各プレイヤーが前回対戦相手が選択した手に応じて今回の手を選択するとしよう。このとき対戦相手の前回選択した手は協調と裏切りの二通りになり、それに対する手は二通りあるから、4通りの戦略が存在する。

戦略の種類を遺伝的アルゴリズム⁹⁾にしたがって、遺伝子タイプとして表現することにしよう。1を協調、0を裏切りとすると、メモリーサイズが1なら、

7) 進化ゲームや「進化的に安定な戦略」をはじめて考え出したのは、Maynard Smith-Price [1973] である。また進化ゲームの全体像については、Maynard Smith [1982] を参照してもらいたい。

8) メモリーの定義については、論者によって異なっている。Axelrod は、対戦相手が、過去にどのような状況のなかで、その手を選択したのかをメモリーに含め、Lindgren and Nordahl [1993]、[1996] は、過去に対戦者が取った手に加えて自分が取った手についてもメモリーに含めている。

9) 遺伝的アルゴリズムの解説書としては、Goldberg [1989]、伊庭 [1994] などがある。

前回の対戦で相手が選択した手に応じて、00, 01, 10, 11 という四つの戦略が存在する。このうち最初の数字は相手が前回0, すなわち裏切りを選択したときに、その戦略がどの手を選択するか、2番目の数字は前回相手が1, すなわち協調を選択したときに、今回その戦略がどの手を選択するのかが示している。したがって各戦略は、つぎのような戦略と考えることができる。

- 00: ALLD 相手が前回どのような手をとっても裏切る。
- 01: TFT 最初は協調を選択する。2回目からは、相手が前回裏切れれば裏切り、相手が前回協調すれば協調する。この戦略はしっぺ返し (Tit For Tat) 戦略と呼ばれる。
- 10: ATFT 最初は裏切りを選択する。2回目からは、相手が前回裏切れれば協調し、協調すれば裏切る。TFT の逆の戦略である。
- 11: ALIC 相手が前回どのような手をとっても協調する。

この四つの戦略が、ランダムに出会って対戦するとしよう。そして一定の回数、対戦が行われた後に、それぞれの個体が獲得した利得の合計に応じて、次世代の個体数が決まるとしよう。コンピュータ・シミュレーションでは、このときには、どの戦略が生き残っていくのかを検討する。

つぎに、メモリーサイズが2のときには、どうなるだろうか。このとき戦略は、対戦相手が前回、前々回に選択した手に対してどのように反応するかによって、以下のように16種類存在する。

戦略のパターンはより複雑になる。たとえば戦略2の0001は、2回続けて相手が協調したときにこちらも協調する、という戦略である。また戦略8の0111は、相手が2回続けて裏切ったときに、はじめてこちらも裏切るというものである。

メモリーサイズが3になると、戦略の種類は表4のように、00000000から11111111までの256種類になる。

表3 メモリーが2のときの戦略

	(前々回対戦相手がとった手, 前回対戦相手がとった手)			
	0,0	0,1	1,0	1,1
戦略 1: 0000 (ALLD)	0	0	0	0
戦略 2: 0001	0	0	0	1
戦略 3: 0010	0	0	1	0
戦略 4: 0011	0	0	1	1
戦略 5: 0100	0	1	0	0
戦略 6: 0101 (TFT)	0	1	0	1
戦略 7: 0110	0	1	1	0
戦略 8: 0111	0	1	1	1
戦略 9: 1000	1	0	0	0
戦略 10: 1001	1	0	0	1
戦略 11: 1010 (ATFT)	1	0	1	0
戦略 12: 1011	1	0	1	1
戦略 13: 1100	1	1	0	0
戦略 14: 1101	1	1	0	1
戦略 15: 1110	1	1	1	0
戦略 16: 1111 (ALLC)	1	1	1	1

表4 メモリーサイズが3のときの戦略の種類

	(3回前, 2回前, 前回, に相手がとった手)							
	0,0,0	0,0,1	0,1,0	0,1,1	1,0,0	1,0,1	1,1,0	1,1,1
戦略 1	0	0	0	0	0	0	0	0
戦略 2	0	0	0	0	0	0	0	1
...
戦略 256	1	1	1	1	1	1	1	1

4 ノイズ

ノイズとは、プレイヤーがミスをする可能性である。たとえば TFT:01 という戦略は対戦相手が前回0を選択したときには0を、前回1を選択したときには1を選択する戦略であるが、本来0を選択するときに1を選択してしまうことをノイズ（ミス）と呼ぶ。

ノイズが存在するケースについては、どのようなことが新たに生じるのだろうか。ノイズは、ずっと裏切り続ける戦略00のような戦略については、それほど大きな影響を及ぼさない。1回ミスしても影響はその1回にとどまるからである。しかし、TFTのような戦略には大きな影響を及ぼす。たとえば TFT と TFT が対戦しているとする。ノイズが存在しないときには、どちらもずっと協調を採用している。しかし、ここで、一方の TFT が協調Cを選択すべきところで、誤って裏切りDをとるとする。そのときには、以後DとCが繰り返されることになる。

TFT CCC CDCDC

TFT CCC DCDCD

つぎに本来Cをとるべきときに誤ってDを選択すると以後Dが続くし、ここでさらに間違えると DCDC というように裏切りあいになる。したがって、ノイズが存在するときには TFT 戦略の優位性は、小さいものになることが予想される。このようなケースでは、「相手が続けて2回裏切ったときに裏切り返す」という戦略が、ノイズに対しては安定になり、より望ましいかもしれないが、この戦略も、1回おきに協調する相手には負けることになる。

III 突然変異と戦略の進化

遺伝的アルゴリズムでは、遺伝子が突然変異によって変化し、環境への適合度 (fitness) が優れた種（染色体）が生き残っていくと考える。繰り返し囚人のディレンマゲームにおいて、たとえば 00000000 というメモリーサイズが3の戦略が存在するとしよう。この戦略は過去にどのような手を相手がとったと

しても、裏切る (0) という手を選択する裏切り者の戦略である。このうちたとえば、3番目の遺伝子が0から1へ突然変異したとする。そのとき遺伝子形は00100000となる。これは、ゲームの対戦相手が過去3回の対戦でとった手が、それぞれ0, 1, 0のときだけ相手と協調し、他のケースでは相手を裏切るという戦略である。

このようにして、新しい戦略が誕生し、そのうち環境に適合している程度が高い戦略が生き残っていくと考える。それが戦略の進化である。メモリーサイズが3のケースでは、潜在的には256種類の戦略が考えられる。

以下で行うコンピュータ・シミュレーションでは、相手が前回取った手にしか反応しない、すなわちメモリーサイズが「実質的に」1の4種類の戦略から出発する。そして、各遺伝子について、ある確率で突然変異が生じるとする。このような突然変異によって、つぎつぎに新しい戦略が生まれ、そのうち高得点をあげた戦略がその個体数を増やしていく。そして、結果として、どのような戦略が支配的になるのかを調べていく。

IV コンピュータ・シミュレーション

1 コンピュータ・シミュレーションの方法

コンピュータ・シミュレーションはつぎのようにして行われる。ただし、以下の説明はメモリーサイズが3のケースについてのものである。メモリーサイズが1と2のケースについては、3のケースから類推してもらいたい。

- ① メモリーサイズが3となる256種類の戦略のうち、4種類の戦略（遺伝子タイプ）戦略1：00000000，戦略86：01010101，戦略171：10101010，戦略256：11111111について初期の個体数を与えてやる。これらの4種類の戦略は遺伝子タイプとしては、8ビットのメモリーサイズが3の戦略であるが、前回相手がとった手にしか反応しないから、そのメモリーサイズは実質的には1である。戦略1は裏切り者、戦略86はTFT，戦略171はATFT，戦略256は正直者にあたる。他の戦略については、初期には存在

しないとする。

- ② 4種類の戦略をとるすべての個体について、対戦以前の3回分のヒストリーを便宜的に与えてやる¹⁰⁾。ヒストリーは8種類存在し、つぎの表5のように分類する。戦略1, すなわち裏切り者については、すべての個体が、1回前の対戦でとった手を0(裏切り), 2回前の対戦でとった手を0, 3回前の対戦でとった手を0とする。また戦略256, すなわち正直者については、1回前が1, 2回前が1, 3回前が1とする。それ以外の戦略86と戦略171については、各ヒストリータイプが1/8ずつだとする。

表5 初期の「ヒストリー」タイプ

	3 回 前	2 回 前	前 回
ヒストリータイプ1	0	0	0
ヒストリータイプ2	0	0	1
ヒストリータイプ3	0	1	0
ヒストリータイプ4	0	1	1
ヒストリータイプ5	1	0	0
ヒストリータイプ6	1	0	1
ヒストリータイプ7	1	1	0
ヒストリータイプ8	1	1	1

- ③ 第1回めの対戦相手となる個体のペア(対戦相手)を乱数で決定する。
第1回めの対戦で、各戦略(遺伝子タイプ)は、相手のヒストリーによって、自分のとる手を選択する。ただし、ノイズが存在するときには、手をミスする可能性が1%存在すると仮定する。
- ④ 第1回めの対戦が行われ、それぞれの利得が決定される。利得行列は表2あるいは表3のように仮定する。第1回の対戦で、各個体を選択した手

10) もちろん対戦前にヒストリーは存在しないわけだから、ここでのヒストリーの与え方は恣意的なものである。このようにしないと、はじめの2回の対戦について、各戦略(遺伝子タイプ)のとる手を決定できないからである。このような方法をとるかわりに、各戦略について、はじめの2回の対戦について、とる手をあらかじめ与えてやる方法もある。

は、その個体の1期前のヒストリーになる。そして、1期前のヒストリーは2期前のヒストリー、2期前のヒストリーは3期前のヒストリーになる。

- ⑤ このようにして10回の対戦が行われ、各個体についての利得の合計を戦略（遺伝子タイプ）ごとに集計し、個体数で割って各戦略の得点（利得）の平均値を求める。また総平均も求める。得点の平均が総平均より大きい戦略（遺伝子タイプ）は、その個体数とその差に比例して増加し、逆に小さい戦略はその差に比例して減少する。比例係数は1とする¹¹⁾。
- ⑥ 各個体のすべての遺伝子について複製を行うが、そのときに0.1%の確率で突然変異が生じると仮定する。すなわち0の遺伝子は0.1%の確率で1に、1の遺伝子は0.1%の確率で0になるとする。
- ⑦ 最後の3回の対戦で、各戦略（遺伝子タイプ）がとった手のパターンをヒストリータイプ1から8に分類し、それを戦略ごとに集計する。その比率にしたがって、つぎの個体のヒストリータイプを決定する。
- ⑧ ③にもどって、対戦を行う。③～⑥のプロセスを200回反復して、各遺伝子タイプの増減を調べる。

2 メモリーサイズが1のケース

メモリー1の戦略は4種類しか存在しない。したがって、はじめから4種類の戦略が存在すると仮定するときには、突然変異によっても新しい戦略が出現することはないから、その影響も存在しないように思える。しかし、実際にはそれほど単純ではない。対戦によって死滅した戦略が、新たに突然変異によって復活する可能性が生まれるからである。

各戦略の優位は、その個体数によって変化する。ある個体数の状況では優位だった戦略が他の状況では不利になることがある。たとえば裏切り者 ALLD (00) という戦略は、正直者 11 が少なく TFT が多い状況では、高得点をあげ

11) 各戦略（遺伝子タイプ）の個体数がマイナスにならないためには、比例係数は1を超えることはできない。

ることができず、やがて死滅するが、正直者が支配的な状況になって、新たに出現すると正直者を騙すことによって個体数を増やすことができるからである。

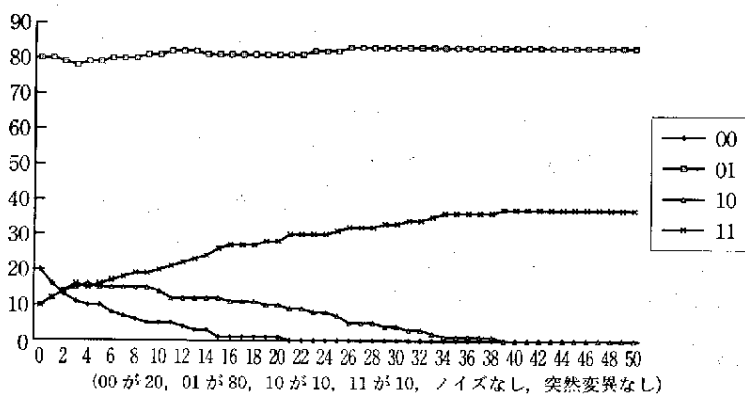
シミュレーション1 (初期の個体数: 00 が 20, 01 が 80, 10 が 10, 11 が 10, ノイズなし, 突然変異なし)

はじめに、ノイズと突然変異が存在しないケースについてシミュレーションを行おう。ここで注意しておくべきなのは、裏切り者 (00) は正直者からは必ず5点をあげることができるということである。また ATFT は前回0を選択している裏切り者と対戦するときには必ず1を選択するから正直者と同様に裏切り者に対しては得点することができない。一方、TFT は、前回0を選択した裏切り者に対しては必ず0を選択するから、どちらも1点しか得点できない。

したがって、正直者や ATFT の個体数が小さいときには、裏切り者は、正直者を食物にできないから、高得点をあげることができず滅亡していきことが予想できる。実際、TFT の個体数を80と大きくし、ALLD を20, ALLC と ATFT を10に押さえたシミュレーション1では、00はすぐに滅亡し、TFT と正直者だけが生き残る結果となった。

ここで、正直者11が裏切り者に勝つのは、どのようなケースなのかをもう

図1 シミュレーション1



少し詳細に検討しておこう。00, 01, 10, 11 の各戦略の総個体数に対する比率をそれぞれ p^{00} , p^{01} , p^{10} , p^{11} とする。このとき戦略 11 の得点の期待値は $3(p^{11}+p^{10})$ である。一方、戦略 00 の得点の期待値は

$$5(p^{11}+p^{10})+(p^{00}+p^{01})$$

である。したがって、11 が 00 に勝つためには、

$$3(p^{11}+p^{01}) > 5(p^{11}+p^{10})+(p^{00}+p^{01})$$

という条件が満たされなければならない。この式を変形すると、

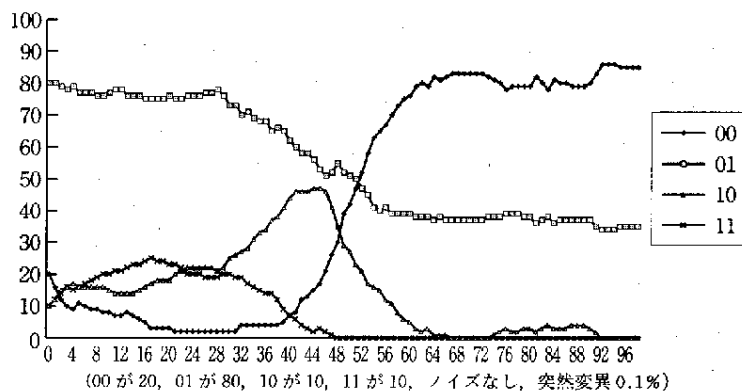
$$2p^{01} > 2p^{11}+5p^{10}+p^{00}$$

となる。すなわち、戦略 01 の比率が十分に高いときには、戦略 11 は戦略 00 に勝つことがわかる。

シミュレーション 2 (初期の個体数: 00 が 20, 01 が 80, 10 が 10, 11 が 10, ノイズなし, 突然変異各遺伝子について 0.1%)

つぎに、このシミュレーション 1 に突然変異を導入してみよう。このときには、一度減んだ 00 が突然変異によって復活することができる。00 は、初期には 10 と 11 の個体数が少ないため一度は滅亡する。しかし、01 と 10, 11 が生

図 2 シミュレーション 2



き残っていて、かつ10と11に個体数が多い状態に00が侵入するのは極めて容易である。このシミュレーションでは、40回あたりから、10と11に勝って、01が急速に個体数を増やしていき、やがて00と01だけが生き残ることになる。

この状態では、00も01もDだけを選択している。このシミュレーション2では、この状態に10や11が侵入することは不可能である。なぜなら、11が00に勝つためには、

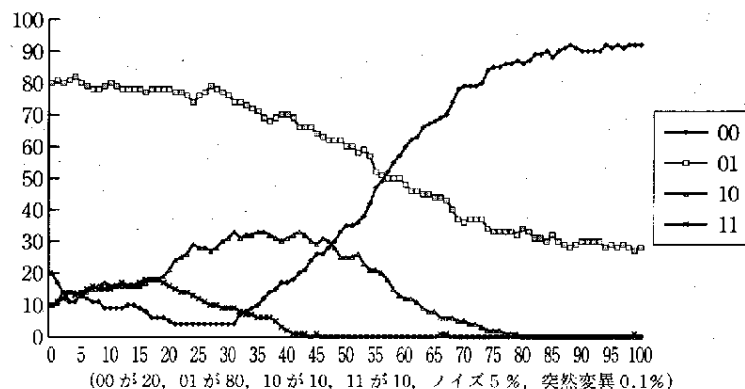
$$3p^{01} > 1$$

という条件が満たされていなければならないが、このケースでは01の個体数は総個体数の1/3以下だからである。

シミュレーション3 (初期の個体数: 00が20, 01が80, 10が10, 11が10, ノイズ各戦略について5%, 突然変異各遺伝子について0.5%)

つぎに、このシミュレーションにおいてノイズがどのような影響をもつのかを調べてみよう。ノイズの効果は複雑であるが、2.4で簡単に考察したように、TFTに対して不利に働くことが予想できる。ノイズを5%として行ったコンピュータ・シミュレーションの結果は、図3のように、図2に比べて、ALLD

図3 シミュレーション3



戦略 (00) がより個体数が増え、TFT (01) は個体数が減少している。

3 メモリーサイズが2のとき

つぎにメモリーサイズが2のケースについてシミュレーションを行ってみた。

シミュレーション4 (初期個体数: 00 が 12, 01 が 80, 10 が 8, 11 が 8, ノイズ 1 %, 突然変異の確率 0.1 %)

このとき、ALLD は、ALLC、ATFT が少ないため、20回めに滅亡してしまった。そのときの戦略の種類と個体数はつぎのようになった。

0001	2	0100	1	0101	79	0111	4
1010	2	1101	3	1111	17		

しかし、対戦回数が増えるにしたがって、新しい戦略が突然変異によって生まれ、何回繰り返しても個体数は収束しなかった。1000回の対戦が終了した時点では、存在する戦略と個体数は、つぎのようになった。

0000	1	0001	3	0011	31	0100	12
0101	27	0110	1	0111	2	1000	10
1001	3	1010	2	1100	10	1101	5
1110	1						

シミュレーション5 (初期個体数: 00, 01, 10, 11 が各 40, ノイズ 1 %, 突然変異の確率 0.1 %)

00, 10, 11 の数が十分多いため、このときには、初期には ALLD が個体数を増やしていき、ほぼ支配的になる。11回めには、0000 の個体数は133個になる。

0000	133	0001	1	0101	25	0111	1
------	-----	------	---	------	----	------	---

しかし、その後新しい戦略が突然変異によって誕生し、ALLDの支配を崩していく。その後は、0110が次第にその個体数を増やしていき、1000回めには、つぎのような状態になる。

0011	3	0100	2	0110	148	0111	1	1100	6
------	---	------	---	------	-----	------	---	------	---

4 メモリーサイズが3のとき

メモリーサイズが3のときには、可能な戦略のタイプは126種類になる。初期の各戦略の個体数の与え方によって、戦略の個体数は変化していくが、いずれにしろ、戦略の種類は収束せず、多様な戦略が共存する状態となった。

シミュレーション6 (初期の個体数: 00000000, 01010101, 10101010, 11111111 が各40個, ノイズ10%, 突然変異の確率0.1%)

00, 01, 10, 11 が各40個の状態から出発する。はじめは、00が11, 10に勝ち、その個体数を増やしていき、01もノイズが存在するので17回めには、

00000000	129	00000001	7	00000010	2	00001000	1
00010000	2	00100000	3	01000000	4	01000101	1
01010101	10	10000000	1				

の状態になる。しかし、突然変異によって産まれた新しい戦略が次第に00に勝利し、1000回目には、

00000000	21	00000001	6	00000010	9	00000011	7
00000100	33	00000101	18	00000110	3	00000111	3
00001000	2	00001100	1	00010001	1	00010010	7
00010100	3	00010101	12	00010111	4	00011101	1
00100000	2	00100101	3	00100110	5	01000000	1
01000010	2	01000100	2	01010000	7	01010001	3

01010010 3 01010101 1

というように多様な戦略が共存した状態になっている。一番個体数が多くなっているのは、00000100 というタイプの戦略であるが、この戦略が他に比べて絶対的に強いわけではなく、戦略の個体数はどんどん変化を続けていく。

シミュレーション 7 (初期の個体数: 00000000 が 16 個, 01010101 が 80 個, 10101010 が 8 個, 11111111 が 8 個, ノイズ 10%, 突然変異の確率 0.1%)

このときには、はじめには、協調型の戦略が勝利する。しかし、新しい戦略が突然変異によって生まれてくるにしたがって、11111111 はその個体数を減少させていく。1000 回の時点では、シミュレーション 6 と同じように、多様な戦略が共存する状態になった。

00000110	1	00001101	2	00010110	1	00100100	1
00100101	1	00101011	1	00101100	2	00101101	1
00101110	1	00101111	11	00111011	2	00111111	1
01000010	1	01000101	14	01000110	22	01000111	2
01001011	11	01001110	1	01001111	2	01010001	1
01010101	4	01010110	1	01011111	2	01100000	1
01100100	2	01100101	6	01100110	6	01101011	1
01110100	1	01110101	2	11000011	1	11000101	1
11000111	1	11010111	1	11100011	1	11100100	1
11101111	1						

V お わ り に

このシミュレーションから、得られた結果を簡単にまとめておこう。まずメモリーサイズが 1 で突然変異、ノイズが存在しないケースについては、裏切り者 00, ATFT10, 正直者 11 の個体数が TFT に比べて大きいときには、裏切

り者が勝ち、TFT の個体数が十分に大きいときには、TFT と正直者が勝利した。しかし、突然変異を導入することによって、正直者の個体数が十分多いところに裏切り者が復活すると裏切り者が勝利する結果となった。そして、ひとたび裏切り者と TFT による支配が成立すると、正直者や ATFT は侵入することができなかった。

メモリーサイズ2と3のシミュレーションについては、確定した結論を得ることはやさしくない。確かに進化の存在は、多様な戦略の共存可能性を示しているように見える。多くのケースについて戦略の種類、個体数は収束せず、最終的に多様な戦略が共存し、その個体数も変動することになった。ただし、どのようなときに、そうなるのかについては、コンピュータ・シミュレーションをより多くのケースについて行い、また理論的に厳密に戦略の優劣を検討する必要がある。そのような研究は現時点では、まだ十分に行われていない。

最後に、この論文の問題点と今後の研究の方向性について少し検討してみた。この論文における進化とは突然変異によって遺伝子が増え、新しい戦略が誕生することであった。しかし、一般には、メモリーサイズ自体が増えるタイプの進化もあるだろう。また遺伝的アルゴリズムで研究されている遺伝子の交差などの可能性も考慮に入れる必要がある。さらにメモリーサイズの定義についても、その大きさを増やすとともに、Axelrod や Lindgren and Nordahl などの定義を取り入れる必要があるかもしれない。このような点について、シミュレーションを改良し、どのようなケースについて協調的な状況が生まれるのかをさらに検討していきたい。

参考文献

- Axelrod, R. [1980a] "Effective Choice in the Prisoner's Dilemma," *Journal of Conflict Resolution*, 24, pp. 3-25.
—— [1980b] "More Effective Choice in the Prisoner's Dilemma," *Journal of Conflict Resolution*, 24, pp. 379-403.
—— [1984] *The Evolution of Cooperation*, Basic Books. (松田裕之訳 [つきあ

い方の科学：バクテリアから国際関係まで」 HBJ 出版局)。

- [1987] "The Evolution of Strategies in the Iterated Prisoner's Dilemma," in *Genetic Algorithms and Simulated Annealing*, ed. by L. Davis, Morgan Kaufmann.
- [1997] *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton University Press.
- Binmore, K. [1992] *Fun and Games: A Text on Game Theory*, D. C. Heath and Company.
- Fudenberg, D. and J. Tirole [1992] *Game Theory*, The MIT Press.
- Goldberg, D. E. [1989] *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley.
- 伊庭斉志 [1994] 『遺伝的アルゴリズムの基礎：GA の謎を解く』 オーム社。
- Kreps, D. M. [1990] *Game Theory and Economic Modeling*, Clarendon Press Oxford.
- Langton, C. G. ed. [1989] *Artificial Life*, Addison-Wesley Publishing Company.
- [1993] *Artificial Life III*, Addison-Wesley Publishing Company.
- [1996] *Artificial Life: An Overview*, The MIT Press.
- Lindgren, K. [1991] "Evolutionary Phenomena in Simple Dynamics" in *Artificial Life II*, eds. by C. G. Langton, J. D. Farmer, S. Rasumussen and C. Taylor, Addison Wesley Publishing Company.
- Lindgren, K. and M. G. Nordahl [1993] "Artificial Food Webs" in *Artificial Life III*, ed. by Langton, C. G., Addison-Wesley Publishing Company.
- Lindgren, K. and M. G. Nordahl [1996] "Cooperation and Community Structure in Artificial Ecosystems" in *Artificial Life: An Overview*, ed. by Langton, C. G., The MIT Press.
- Maynard Smith, J. [1982] *Evolution and the Theory of Games*, Cambridge University Press. (寺本英・梯正之訳『進化とゲーム理論：闘争の論理』産業図書, 1985年)。
- Maynard Smith, J. and G. A. Price [1973] "The Logic of Animal Conflict," *Nature*, 246, pp. 15-18.
- Poundstone, W. [1992] *Prisoner's Dilemma*, Doubleday. (松浦俊輔他訳『四人のディレンマ』青土社, 1995年)。
- Robinson, J. [1971] *Economic Heresies*, London, Macmillan. (宇沢弘文訳『異端の経済学』日本経済新聞社, 1973年)。
- Sraffa, P. [1960] *Production of Commodities by Means of Commodities*, Cambridge,

Cambridge University Press. (菱山泉・山下博訳『商品による商品の生産：経済学批判序説』有斐閣, 1962年)。

高増明・服部純典 [1999a] 「協調の進化：繰り返し囚人のディレンマ・ゲームのコンピュータ・シミュレーション」『大阪産業大学論集 社会科学編』111号, 55-68ページ。

高増明・服部純典 [1999b] 「繰り返し囚人のディレンマ・ゲームにおけるノイズとメモリーサイズ」『大阪産業大学論集 社会科学編』113号, 91-103ページ。